

<https://helda.helsinki.fi>

---

## Supporting Self-Regulated Learning with Visualizations in Online Learning Environments

Ilves, Kalle

ACM

2018-02-21

---

Ilves , K , Leinonen , J & Hellas , A 2018 , Supporting Self-Regulated Learning with Visualizations in Online Learning Environments . in SIGCSE '18: Proceedings of the 49th ACM Technical Symposium on Computer Science Education . ACM , New York , pp. 257-262 , ACM Technical Symposium on Computer Science Education , Baltimore , Maryland , United States , 21/02/2018 . <https://doi.org/10.1145/3159450.3159509>

---

<http://hdl.handle.net/10138/315997>

<https://doi.org/10.1145/3159450.3159509>

---

acceptedVersion

---

*Downloaded from Helda, University of Helsinki institutional repository.*

*This is an electronic reprint of the original article.*

*This reprint may differ from the original in pagination and typographic detail.*

*Please cite the original version.*

# Supporting Self-Regulated Learning with Visualizations in Online Learning Environments

Kalle Ilves, Juho Leinonen, Arto Hellas  
University of Helsinki  
Helsinki, Finland  
{firstname.lastname}@helsinki.fi

## ABSTRACT

In this article, we study how visualizations could be used to support students' self-regulation in online learning. We conducted a randomized controlled trial with three groups: one control group without visualization, one treatment group with textual visualization, and one treatment with graphical visualization with information on peers' average achievement. We studied how different visualizations affect students' academic performance and behavior. We focused on four factors; starting, scheduling, earliness and exercise points, where the first three are related to time management and self-regulation. The last factor measures course performance in terms of completed exercises. Our results suggest that the lowest performing students can benefit from a visualization, whereas the highest performing students are not affected by the presence or absence of a visualization. We also found that visualizations that do not provide the means to compare your own performance with others may even be harmful to performance oriented students.

## CCS CONCEPTS

•Human-centered computing → Visualization; Visualization systems and tools; •Applied computing → Education; Interactive learning environments; E-learning;

### ACM Reference format:

Kalle Ilves, Juho Leinonen, Arto Hellas. 2018. Supporting Self-Regulated Learning with Visualizations in Online Learning Environments. In *Proceedings of The 49th ACM Technical Symposium on Computing Science Education, Baltimore, MD, USA, February 21–24, 2018 (SIGCSE '18)*, 6 pages. DOI: 10.1145/3159450.3159509

## 1 INTRODUCTION

Students have been found to have different motivations for trying to achieve learning goals [10]; in broad terms, students can be categorized as mastery oriented learners who seek to understand and develop their competence, and as performance oriented learners who seek to perform well and contrast their performance to others. Our study focuses both on visualizations for supporting self-regulated learning as well as on students' motivations.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).  
SIGCSE '18, Baltimore, MD, USA

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.  
978-1-4503-5103-4/18/02...\$15.00  
DOI: 10.1145/3159450.3159509

We examine how different visualizations affect students' academic performance and behavior in an online learning environment. For the purposes of this study, we developed two different visualizations: a textual and a radar visualization. These visualizations show the students their exercise points and time management when completing the problems on a course. The textual visualization gave the students a score on each measured factor and provided textual feedback on how to improve their score, while the radar visualization also provided the means to compare the performance to others in the class. The two visualizations were developed with students' different motivations in mind: by comparing the visualizations' impact on performance of differently motivated students we could potentially find a way to support differently motivated students with different kinds of visualizations.

Academic performance was measured through the number of exercise points, which was a convenient measure for our course and an often used measure in previous studies. We also measured whether the visualizations can be used to influence how early and on how many days the students solve exercises before the deadline to see whether a visualization encourages behavior which has been shown to lead to better academic performance. These behavioral measures on top of the number of exercise points were the key visualized aspects in the visualizations we developed. As a summary, our overarching research question for this study is “How do visualizations that support self-regulated learning affect students' academic performance and behavior?”

Achievement goals and visualizations have been previously studied by, e.g., Auvinen [3], who suggested that students' performance is only slightly influenced by visualizations. Achievement goals have been also studied by Zingaro and Porter [17, 18], who suggested that both performance and mastery orientation can influence students grades and exam scores. Similarly, students' behavior has been studied by e.g. [12], who sought for interventions that reduce procrastination.

This work is organized as follows. Section 2 covers the background information related to scheduling and completing programming exercises, goal orientation in self-regulated learning, and visualizations developed for supporting self-regulated learning. Section 3 addresses the methodology of our study, describing the course in which this study was conducted and its learning environment, the visualizations that were used, and other key aspects of the study. In Section 4 we present and discuss the results of our study. Finally, Section 5 concludes this work by summarizing our findings and outlining the future work that we are planning to conduct.

## 2 BACKGROUND

### 2.1 Scheduling and completing exercises

Students' academic performance on a course has been measured using for example factors like the amount and quality of completed exercises, exam scores, the final grade, and drop-out-rate [5, 11, 13, 16]. Studies have revealed that the total number of exercises the student has completed, the number of days that the student has been working on the exercises and how early the student starts working on the exercises correlate with students' performance.

Spacco et al. [16] studied the relationship between the amount of completed exercises and students' exam scores. In the study, students completed a set of programming exercises which were similar to the programming questions asked in the exam. The study revealed that the amount of exercises attempted and answered had a positive correlation with exam scores.

Spacing out learning over time has been shown to be a more effective way of learning compared to learning the same material in a single session [5]. Leppänen et al. [11] studied how students space out their work in a programming course. Students participating in the study attended school work on average three times a week. The study revealed that students with higher exam scores averaged less working days per week, while the study conducted by Cepeda et al. [4] indicates that spacing out work over time is beneficial for students' performance.

Martin et al. [13] studied how starting school work early or starting it late affected the quality of finished work. The study was conducted on a course which involved four separate programming projects, with students being given approximately one month for each assignment. The study revealed that individual students underestimated the effort required to complete the assignments 72% of the time.

### 2.2 Goal orientation in self-regulated learning

Goal orientations can be defined as purposes for achievement behavior [10]. Rather than focusing on *what* people try to achieve, goal orientation characterizes the individual's *motivation* for achieving various objectives. Adopting different orientations is associated with different quality of engagement in school work as well as with different emotional experiences in school.

Perception of the number and the definition of orientations that people may adopt in achievement situations vary among researchers. Nevertheless, goal orientation research focuses mostly on two main orientations: *mastery* and *performance* goals. Mastery orientation refers to the purpose of developing competence. Mastery-oriented students focus on learning, understanding, developing skills, and mastering information. Performance orientation on the other hand refers to the purpose of demonstrating competence. Performance-oriented students strive to perform well relative to others and seek to appear competent.

Mastery and performance orientation can be divided into *approach* and *avoidance* [6]. When pursuing performance-approach goals, the student is oriented towards demonstrating high ability. In contrast, a student pursuing performance-avoidance goals is seeking to avoid showing low competence. Similarly, students with a mastery-approach seek to master the task, while the mastery-avoidance students seek to avoid failing in the task.

Mastery-oriented students have been suggested to also have higher self-efficacy, persistence, preference for challenge, self-regulated learning and positive affect and well-being [10]. Performance-oriented students, on the other hand, have been found to be associated with positive outcomes such as persistence and positive affect as well as negative outcomes such as anxiety, disruptive behavior and low retention of knowledge [10].

It is common to assess the existence of mastery and performance goal orientations with questionnaires constructed with inventories using Likert type scales [10]. In this study, we use the questionnaire by Elliot and McGregor [7].

### 2.3 Visualizations for supporting self-regulated learning

Multitude of systems supporting students' monitoring and reflection of their own behavior have been developed. The impact of such systems to students' performance is typically evaluated both empirically and using feedback surveys or usability studies. The following systems share similarities with our study in that they also visualize students' performance with different variables such as exercise points and time spent in the learning materials, and also provided a way to compare own performance to peers' performance.

*iTree* [14] is a mobile based system, which allows students to interact, exchange information, engage in discussion, and collaborate on projects. It visualizes the activity of each individual student as a tree that changes its form depending on the student's number of posts, number of times posts are read, number of replies to posts and the ratio of total forum posts to replies. Students who used the system spent more time reading posts and created more posts on the discussion forum than the students who did not use it.

*Student Activity Meter (SAM)* [8] aims to assist both teachers and learners with self-reflection and awareness of what and how students are doing. It visualizes, e.g., work periods on a time line, number of accessed resources and the time spent with them, forum activities and comparison to peers. The system also recommends resources based on the resources accessed by the student.

*StepUp!* [15] empowers students to reflect on their own activity, and that of their peers, in open learning environments. The system tracks the time students spend in different activities on websites and the Eclipse development environment. It visualizes social interaction, time spent, artifacts produced, resource use, and allows comparison to other students. Feedback received from students suggested that the system helps students to assess how they are doing on the course and allows them to better compare themselves with their peers. Many students considered the feeling of being observed as a negative aspect.

*Course Signals* [1] allows teachers to send personalized emails to students that contain information about their current performance in a given course, encourage students to visit various help resources on campus or office hours and employs learning analytics to predict students' performance. The performance is measured using past academic history and the amount of interaction with the learning environment. At-risk students are sent emails that visualize their risk level as a traffic light and include pointers to various help resources on the campus. A study revealed that courses that implement Course Signals had a strong increase in satisfactory grades, and a decrease in unsatisfactory grades and withdrawals.

Finally, Auvinen et al. [3] developed a set of eight different visualizations visualizing students' average number of attempts per exercise, average points of the first attempts to each exercise, average time interval between consecutive attempts to the same exercise and average distance of the submissions to the deadline. Six of the visualizations are heat maps where color represents past students' final exercise points based on their behavior. Two of the visualizations are line charts with a thick line representing final exercise points and a gray area lower and upper quartiles. All the visualizations have a "You are here" label plotting student's own behavior, allowing students' to compare themselves to peers. The majority of students found visualizations reasonably easy to interpret but claimed that visualizations did not affect how exercise were done and submitted. There was also no significant change observed in the earned exercise points from the group who used the visualizations.

### 3 METHODOLOGY

#### 3.1 Context and Data

The data for this study comes from an eight and a half-week online introductory programming course that was organized during the Spring 2017 at University of Helsinki, a European research first University. The course uses a blended online textbook with videos, instructional materials and embedded programming assignments that were automatically assessed. The course had a total of seven exercise sets that were released weekly; at the beginning of the course, the participants had a 20-day window to the first deadline. There was no limit or penalty for submitting incorrect answers, and the students could resubmit the assignments as many times as they wished.

The course had a total of 753 students, from whom 442 were included in this study. The 311 students were excluded for reasons including not wanting to participate in the study, not answering the provided questionnaires, or not attending more than a single week of the course.

The 442 students were divided randomly into three groups. The control group consisted of 168 students, the treatment group A consisted of 141 students, and the treatment group B consisted of 133 students. For analysis, the three groups were then divided further into four groups based on their goal orientation determined by the survey by Elliot and McGregor [7]. In both treatment groups and the control group, the vast majority of the students were classified as mastery approach students. A small smoothing for orientations with less students was applied to distribute students with weaker mastery approach orientation to the second strongest orientation. The final group sizes used for the analysis in this study are shown in Table 1.

In this study, we focus on the overall populations in the different conditions as well as on the influence of the treatments to performance and mastery approach oriented students.

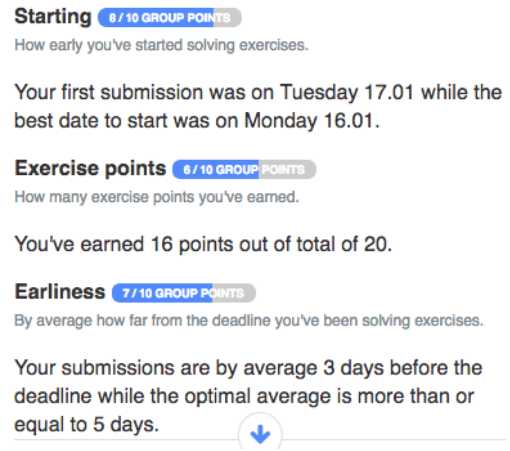
#### 3.2 Treatment and Visualizations

We attempted to influence students' working habits using visualizations embedded into the course materials. The specific metrics that we chose to visualize were chosen based on earlier work indicating that they influence students' academic performance [5, 11, 13, 16].

**Table 1: Sizes of the control group and the treatment groups which are divided into goal orientations.**

	Control	Grp A/Radar	Grp B/Text
mastery approach	55	45	48
performance approach	49	49	42
mastery avoidance	35	27	25
performance avoidance	29	20	18
total	187	141	133

**Figure 1: The textual visualization of a problem set. The students could see the Scheduling metric by scrolling down.**

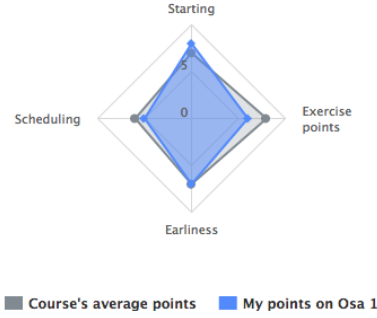


The four metrics that were used to describe students' behavior are as follows: *starting*, *scheduling*, *earliness* and *exercise points*. *Starting* measured how soon after the release of an exercise set the student started to work on the exercises; *Scheduling* measured the number of days that the student worked on an exercise set; *Earliness* measured the average distance from the deadline of a specific exercise set; and *Exercise points* measured how many points the student had received from the exercises of a certain exercise set. Each metric was normalized to a range from 0 to 10, where 0 indicated poorest performance and 10 indicated the best performance. The visualizations showed data for each separate week in the course, which corresponded to separate pages in the course material.

Two types of visualizations were used: the *textual* and the *radar* visualization. The textual visualization showed the variables as a list of values. Each variable had the name of the metric, a short description of it, the value of the metric as a progress bar with a numerical label and a description of where the points came from and how to maximize the value of the metric (see Figure 1). For example, for the scheduling metric, the student could see a guiding text "You've been solving exercises on 2 days while the optimal number of days is at least 3".

The radar visualization showed the metrics using a radar chart. The radar chart consisted of a sequence of spokes where each spoke represented one of the four metrics. The length of each spoke was defined by the value of the metric the spoke was representing. The

Figure 2: The radar visualization of a problem set.



same textual suggestions than in the textual visualization were available when hovering on the values using a mouse. The radar chart was layered so that the student's performance was displayed using a blue layer and the average performance of all students in the course was displayed as a gray layer (see Figure 2).

The differences in the visualizations were motivated by performance oriented students striving to perform well relative to others whereas mastery oriented students strive to develop competence [10]. In the radar visualization, students had an opportunity to compare their points to their peers' points which should motivate performance oriented students. The textual visualization on the other hand was more focused in visualizing the student's own behavior and provided suggestions on how to improve it. We hypothesized that this would motivate mastery oriented students.

Group A was assigned the radar visualization, and Group B was assigned the textual visualization.

## 4 RESULTS AND DISCUSSION

Table 2 lists the means (denoted by  $\mu$ ) and standard deviations (denoted by  $\sigma$ ) of the variable values for control and treatment groups and for goal orientations within the groups. The size of each group is denoted by  $n$ .

### 4.1 Visualizations and exercise completion

When comparing the average exercise points from the control group ( $\approx 160.35$ ) to the treatment groups ( $\approx 170.23$  for radar and  $\approx 147.67$  for textual), no statistically significant difference was found. This applies to comparing the treatment groups as a whole to the control group as well as comparing both treatment groups individually to the control group.

When comparing the treatment groups, Wilcoxon test revealed a statistically significant difference between the treatment groups ( $p \approx 0.014$ ). Students in the radar visualization group earned on average 170.23 exercise points, whereas students in the textual visualization group earned on average 147.67 exercise points. The difference between the treatment groups is approximately 22.56 exercise points. In other words, students in the radar group completed approximately 15% more exercises than the students in the textual group.

We then examine the control group and treatment groups further by dividing them into tertiles based on the number of exercise

Table 2: Mean and standard deviation values for exercise points and behavioral variables for the different groups in the study.

Control	exercise points	starting	earliness	scheduling
total $n = 187$	$\mu \approx 160.35$ $\sigma \approx 68.94$	$\mu \approx 8.19$ $\sigma \approx 6.92$	$\mu \approx 12.99$ $\sigma \approx 7.38$	$\mu \approx 2.89$ $\sigma \approx 1.43$
mast. app. $n = 55$	$\mu \approx 164.98$ $\sigma \approx 68.82$	$\mu \approx 9.1$ $\sigma \approx 7.11$	$\mu \approx 12.51$ $\sigma \approx 7.69$	$\mu \approx 2.88$ $\sigma \approx 1.53$
perf. app. $n = 49$	$\mu \approx 173.59$ $\sigma \approx 65.9$	$\mu \approx 8.3$ $\sigma \approx 6.7$	$\mu \approx 13.86$ $\sigma \approx 7.48$	$\mu \approx 3.05$ $\sigma \approx 1.42$
mast. avoid. $n = 35$	$\mu \approx 121.77$ $\sigma \approx 72.3$	$\mu \approx 4.38$ $\sigma \approx 3.99$	$\mu \approx 12.84$ $\sigma \approx 7.12$	$\mu \approx 2.62$ $\sigma \approx 1.45$
perf. avoid. $n = 29$	$\mu \approx 145.66$ $\sigma \approx 76.72$	$\mu \approx 9.15$ $\sigma \approx 8.32$	$\mu \approx 10.48$ $\sigma \approx 7.21$	$\mu \approx 2.46$ $\sigma \approx 1.41$
Radar	exercise points	starting	earliness	scheduling
total $n = 141$	$\mu \approx 170.23$ $\sigma \approx 61.42$	$\mu \approx 8.64$ $\sigma \approx 6.55$	$\mu \approx 13.45$ $\sigma \approx 7.18$	$\mu \approx 3.12$ $\sigma \approx 1.55$
mast. app. $n = 45$	$\mu \approx 176.27$ $\sigma \approx 59.43$	$\mu \approx 8.52$ $\sigma \approx 6.2$	$\mu \approx 14$ $\sigma \approx 7.1$	$\mu \approx 3.5$ $\sigma \approx 1.72$
perf. app. $n = 49$	$\mu \approx 185.06$ $\sigma \approx 58.7$	$\mu \approx 9.92$ $\sigma \approx 6.63$	$\mu \approx 13.5$ $\sigma \approx 7.33$	$\mu \approx 2.84$ $\sigma \approx 1.47$
mast. avoid. $n = 27$	$\mu \approx 127.48$ $\sigma \approx 72.14$	$\mu \approx 6.29$ $\sigma \approx 7.19$	$\mu \approx 11.5$ $\sigma \approx 8.01$	$\mu \approx 2.77$ $\sigma \approx 1.54$
perf. avoid. $n = 20$	$\mu \approx 138.15$ $\sigma \approx 70.1$	$\mu \approx 7.8$ $\sigma \approx 6.98$	$\mu \approx 10.98$ $\sigma \approx 7.11$	$\mu \approx 2.6$ $\sigma \approx 1.6$
Textual	exercise points	starting	earliness	scheduling
total $n = 133$	$\mu \approx 147.67$ $\sigma \approx 71.13$	$\mu \approx 8.76$ $\sigma \approx 6.55$	$\mu \approx 11.17$ $\sigma \approx 6.46$	$\mu \approx 2.70$ $\sigma \approx 1.62$
mast. app. $n = 48$	$\mu \approx 158.54$ $\sigma \approx 65.26$	$\mu \approx 8.99$ $\sigma \approx 6.94$	$\mu \approx 11.94$ $\sigma \approx 6.83$	$\mu \approx 3.08$ $\sigma \approx 1.7$
perf. app. $n = 42$	$\mu \approx 155.12$ $\sigma \approx 71.49$	$\mu \approx 10.31$ $\sigma \approx 6.41$	$\mu \approx 10.6$ $\sigma \approx 5.86$	$\mu \approx 2.46$ $\sigma \approx 1.38$
mast. avoid. $n = 25$	$\mu \approx 105.4$ $\sigma \approx 71.05$	$\mu \approx 5.05$ $\sigma \approx 4.63$	$\mu \approx 10.23$ $\sigma \approx 6.45$	$\mu \approx 2.23$ $\sigma \approx 1.4$
perf. avoid. $n = 18$	$\mu \approx 151.78$ $\sigma \approx 69.26$	$\mu \approx 9.22$ $\sigma \approx 6.09$	$\mu \approx 11.13$ $\sigma \approx 6.67$	$\mu \approx 2.79$ $\sigma \approx 1.92$

points earned, we notice that the exercise points vary in the lowest two thirds. Table 3 displays the mean exercise points (denoted by  $\mu$ ) and the standard deviations (denoted by  $\sigma$ ) of the lowest, middle and highest thirds of the treatment groups and the control group. The size of each group is denoted by  $n$ .

The highest performing students earned on average almost all of the available points in the course (the maximum was 216 exercise points) with a standard deviation of less than two in each of the treatment groups and in the control group. This indicates that the highest performing students, regardless of the visualization, earned

**Table 3: Exercise points of the lowest, middle and highest tertiles.**

	lowest third	middle third	highest third
control $n = 164$	$\mu \approx 65.02$ $\sigma \approx 25.83$ $n = 54$	$\mu \approx 194.20$ $\sigma \approx 24.89$ $n = 55$	$\mu \approx 215.89$ $\sigma \approx 0.31$ $n = 55$
radar $n = 137$	$\mu \approx 86.31$ $\sigma \approx 36.42$ $n = 45$	$\mu \approx 204.15$ $\sigma \approx 14.15$ $n = 46$	$\mu \approx 215.87$ $\sigma \approx 0.34$ $n = 46$
textual $n = 132$	$\mu \approx 55.98$ $\sigma \approx 22.18$ $n = 44$	$\mu \approx 171.12$ $\sigma \approx 32.56$ $n = 43$	$\mu \approx 214.91$ $\sigma \approx 1.70$ $n = 45$

the maximum points in the course without nearly any variance. A hypothesis for the reasons behind these results is that the students who have strong task related or self-regulatory skills do not benefit from the external feedback provided by the visualizations as much as the students with weaker skills.

The lowest two thirds reveal a difference between the groups. Students in the middle third are akin except when using the textual visualization, which seems to be harmful. Students in the lowest third of the radar visualization group earned on average 86.31 exercise points, whereas students in the lowest third of the textual visualization group earned on average 55.98 exercise points. Additionally, the lowest third in the control group earned on average 65.02 exercise points. The differences between the radar visualization group and the other groups were statistically significant (Wilcoxon,  $p \approx 0.0029$  for control,  $p < 0.001$  for textual), while the difference between the textual visualization and the control group was not statistically significant. This finding emphasizes that even if the positive impact of a visualization might not be that large on the whole group of students, positive impacts can be observed within student subpopulations.

## 4.2 Visualizations and goal orientation

Exercise points earned by the performance approach oriented students differed significantly in the treatment groups (Wilcoxon,  $p \approx 0.01$ ). Performance approach oriented students in the radar visualization group complete more exercises than the performance approach oriented students in the textual visualization group: students in the radar visualization group earned on average nearly 30 more exercise points than the ones in the textual visualization group. This difference is also in line with goal orientation theory, which suggests that performance approach oriented students draw motivation from outperforming others – something that is supported by the radar visualization. However, at the same time, the mastery oriented students also performed better in the radar group than the students in the textual group.

Performance approach oriented students in the control group earned approximately on average 173.59 exercise points, while they earned on average 155.12 points in the textual visualization group. This difference was confirmed significant by a Wilcoxon test ( $p \approx 0.027$ ). That is, the performance approach oriented students who had no visualization performed better than the performance

approach oriented students with the textual visualization. This indicates that a visualization may even be harmful to performance approach oriented students if the visualization does not provide a comparison to other students.

This observation suggests that performance and mastery approach oriented students may benefit from different types of visualizations – or from the lack of a visualization. To our knowledge, previous work has not reported negative effects of visualizations on students. One reason for the lack of observed negative effects could be that most studies only measure the effects of the visualizations on the whole population of students without diving students by their goal orientation and measuring their differences. However, Auvinen et al. [2] divided the student population by their goal orientation but did not find any negative effects in any student groups. Another possible reason for lack of research on negative effects could be that reporting positive results is far more common than reporting negative results [9].

## 4.3 Visualizations and time management

The length of the delay in days students submit their first exercise after an exercise set is published (*starting*) does not vary between groups with a significant difference. However, the mean difference in days between all the student's exercise submissions during an exercise set and the deadline of the exercise set (*earliness*) differed significantly between the groups. Students in the radar visualization group submit exercises on average 13.45 days before the deadline whereas students in the textual visualization group submit exercises on average 11.17 days before the deadline. A Wilcoxon test suggests that the difference can be statistically significant ( $p \approx 0.04$ ).

Exercise submission earliness of the performance approach oriented students varied in the treatment groups. In the radar visualization group performance approach oriented students submitted exercises approximately on average 13.5 days before the deadline whereas performance approach oriented students in the textual visualization group 10.6 days before the deadline. A Wilcoxon test again confirmed that the groups differed significantly ( $p \approx 0.047$ ). Similar difference in exercise submission earliness was observable between the performance approach oriented students in the control group and the textual visualization group. Performance approach oriented students in the control group submitted exercises 3.26 days earlier than the performance approach oriented students in the textual visualization group. A Wilcoxon test confirmed a significant difference between the groups ( $p \approx 0.038$ ).

When considering scheduling, i.e. dividing the work across multiple days, we observe that the visualizations did not increase the number of days during which the students worked on the assignments. On the contrary, performance approach oriented students in the textual visualization group submitted exercises on average on 2.46 days whereas performance approach oriented students in the control group submitted exercises on average on 3.05 days. The difference between the groups was statistically significant ( $p \approx 0.031$ ). It is possible that the visualized metrics influence each other; the performance approach oriented may have striven to gain all the exercise points as fast as possible, ignoring the feedback related to spacing out their effort.

#### 4.4 Limitations of the study

Next, we address both internal and external validity concerns related to our study. While our visualizations were evaluated using a randomized between-subject experiment which provides a good confidence that the differences between the treatment and control groups are due to the intervention, there is a possibility of random variations in the pre-test performances of the groups. It is also possible that some of the students learned of the experiment from their peers, which could have affected the study outcomes.

It is also possible that there is a self-selection bias in the goal orientation and in the final survey feedback. This might have resulted in some goal orientation or opinions to have a stronger presence than others. For example, avoidance-oriented students might skip some of the survey to minimize effort, or to avoid checking or noticing if they perform poorly when compared to others.

The results could be dependent on the arrangements of the course, the participants of the course, and the culture. Our course was a university level course and the participants provide a very biased sample. In a more competitive culture, the effects related to the performance approach oriented students might also manifest in a stronger fashion. Also, in some cultures students may be more dependent on external support, which can affect how they behave in self-regulated learning activities.

Finally, any research is as valid as the used measurements are. In this study, we used the questionnaire by Elliot and McGregor [7], which may not be optimal for assessing mastery and performance orientation – similarly, the link between different goal orientations and course performance needs to be studied further in different contexts.

#### 5 CONCLUSIONS

In this work, we studied whether different types of visualizations embedded to the course materials can influence students' performance on the visualized variables. We conducted a randomized controlled trial, where the student population of an online programming course was divided into three groups, where one group had no visualizations, one group had a textual visualization, and one group had a radar visualization that also compared their progress to the progress of others in the course. Furthermore, we studied whether students' goal orientation influences their behavior.

Our study suggests that visualizations can have both negative and positive effects on students. Based on our results, the lowest performing students can benefit from a visualization the most, whereas highest performing students completed on average all the exercises on the course regardless of their visualization. Our results also indicate that a visualization that does not provide the means to compare your own performance to the performance of others can even be harmful to performance approach oriented students. The students with the radar visualization had overall better performance than the textual visualization group, and in some cases, the students with the textual visualization actually performed worse than the students in the control group with no visualization.

There was no clear indication that the visualizations would have a positive effect in behavior which is considered to lead into better academic performance. On the contrary, performance approach oriented students in the textual visualization group submitted exercises closer to deadline than those in the control group who did

not receive any visualization. On the other hand, the students in the radar visualization group submitted exercises earlier than students in the textual visualization group. This difference was most significant within the performance approach oriented students.

Our results can be used as a guideline when designing course material visualizations; there is no clear consensus on the best visualization, but it is clear that students' goal orientations should also be taken into account. As a part of our future work, we are conducting a similar controlled experiment that studies only the radar visualization. In the study reported here, the students with the radar visualization performed on average better than the students in the other conditions. We hypothesize that the reason for the improved performance was that the students could compare their behavior to others, but as we did not explicitly study two versions of the radar visualization, this has to be confirmed.

#### REFERENCES

- [1] Kimberly E Arnold and Matthew D Pistilli. 2012. Course signals at Purdue: using learning analytics to increase student success. In *Proc. of the 2nd international conference on learning analytics and knowledge*. ACM, 267–270.
- [2] Tapio Auvinen et al. 2015. Educational Technologies for Supporting Self-Regulated Learning in Online Learning Environments. (2015).
- [3] Tapio Auvinen, Lasse Hakulinen, and Lauri Malmi. 2015. Increasing students' awareness of their behavior in online learning environments with visualizations and achievement badges. *IEEE TLT* 8, 3 (2015), 261–273.
- [4] Nicholas J Cepeda, Harold Pashler, Edward Vul, John T Wixted, and Doug Rohrer. 2006. Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological bulletin* 132, 3 (2006), 354.
- [5] Frank N Dempster. 1988. The spacing effect: A case study in the failure to apply the results of psychological research. *American Psychologist* 43, 8 (1988), 627.
- [6] Andrew J Elliot and Marcy A Church. 1997. A hierarchical model of approach and avoidance achievement motivation. *Journal of personality and social psychology* 72, 1 (1997), 218.
- [7] Andrew J Elliot and Holly A McGregor. 2001. A 2×2 achievement goal framework. *Journal of personality and social psychology* 80, 3 (2001), 501.
- [8] Sten Govaerts, Katrien Verbert, Erik Duval, and Abelardo Pardo. 2012. The student activity meter for awareness and self-reflection. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems*. ACM, 869–884.
- [9] John PA Ioannidis. 2005. Why most published research findings are false. *PLoS med* 2, 8 (2005), e124.
- [10] Avi Kaplan and Martin L Maehr. 2007. The contributions and prospects of goal orientation theory. *Educational Psychology Review* 19, 2 (2007), 141–184.
- [11] Leo Leppänen, Juho Leinonen, and Arto Hellas. 2016. Pauses and spacing in learning to program. In *Proc. of the 16th Koli Calling International Conference on Computing Education Research*. ACM, 41–50.
- [12] Joshua Martin, Stephen H. Edwards, and Clifford A. Shaffer. 2015. The Effects of Procrastination Interventions on Programming Project Success. In *Proc. of the Eleventh Annual International Conference on International Computing Education Research (ICER '15)*. ACM, New York, NY, USA, 3–11. <https://doi.org/10.1145/2787622.2787730>
- [13] Joshua Martin, Stephen H Edwards, and Clifford A Shaffer. 2015. The Effects of Procrastination Interventions on Programming Project Success. In *Proc. of the eleventh annual International Conference on International Computing Education Research*. ACM, 3–11.
- [14] Jun Nakahara, Shinichi Hisamatsu, Kazaru Yaegashi, and Yuhei Yamauchi. 2005. iTree: Does the mobile phone encourage learners to be more involved in collaborative learning?. In *Proc. of the 2005 conference on Computer support for collaborative learning: learning 2005: the next 10 years!* International Society of the Learning Sciences, 470–478.
- [15] Jose Luis Santos, Katrien Verbert, Sten Govaerts, and Erik Duval. 2013. Addressing learner issues with StepUp!: an evaluation. In *Proc. of the Third International Conference on Learning Analytics and Knowledge*. ACM, 14–22.
- [16] Jaime Spacco, Paul Denny, Brad Richards, David Babcock, David Hovemeyer, James Moscola, and Robert Duvall. 2015. Analyzing student work patterns using programming exercise data. In *Proc. of the 46th ACM Technical Symposium on Computer Science Education*. ACM, 18–23.
- [17] Daniel Zingaro. 2015. Examining interest and grades in computer science 1: A study of pedagogy and achievement goals. *ACM Transactions on Computing Education (TOCE)* 15, 3 (2015), 14.
- [18] Daniel Zingaro and Leo Porter. 2016. Impact of student achievement goals on CS1 outcomes. In *Proc. of the 47th ACM Technical Symposium on Computing Science Education*. ACM, 279–296.